

	human	tez1	EST2	p123
AKFLHMLMSVVVVELLSRFFVVTETTFQKNR	ISEIEWLVLGKRSNAKMCLSDFEKRKQIFAETIYWLNSFLIPIILQSFYFVITESSDLNRN	LKDFRWLFISD--IWFTKHFNENQLAICFISWLFRLIPKLIQTFYFCTEISSTVT-	TREISMMQVET-SAKHFYYFDHEN-IYVLWKLRLIFEDLVSLRCFFVYFTEQQKSYSK	**** *

	human	tezi	EST2	p123		Motif 1
	LFYRKSVWSKLSQSIGIRQHLKRVQLRDVSEAEVQRHREAPPALLTSRLRFIPKP--DGL	TVYFRKDIWKLCPFI-TSMKMEAFEKINENNVMDTQK-TTLPPAVIRLLPKK--NTTF	IVYFRHDTWNKLIPTFIVEYFKTYLVENNVCRNHNVSYTLS--NFNHSKMRIIPKSKSNEEF	TYYYRKNIWDVIMKMSI-ADLKETTAEVQKEVEEWKKS-LGFAPGKLRLLIPKK--TTF	*	*
				*	*	*
				*	*	*

human
tezi
EST2
p123

tez1 KKDLLKHRMFGRR-KKYFVVRIDIKSCYDRIKQDLMFRIYVK-KLKDPFVIRKYATHTATS
EST2 KQRLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTIN
p123 EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSITFLKTTTKLLSSDFWIMTAQILKRRKN

FIG. 1

2/34

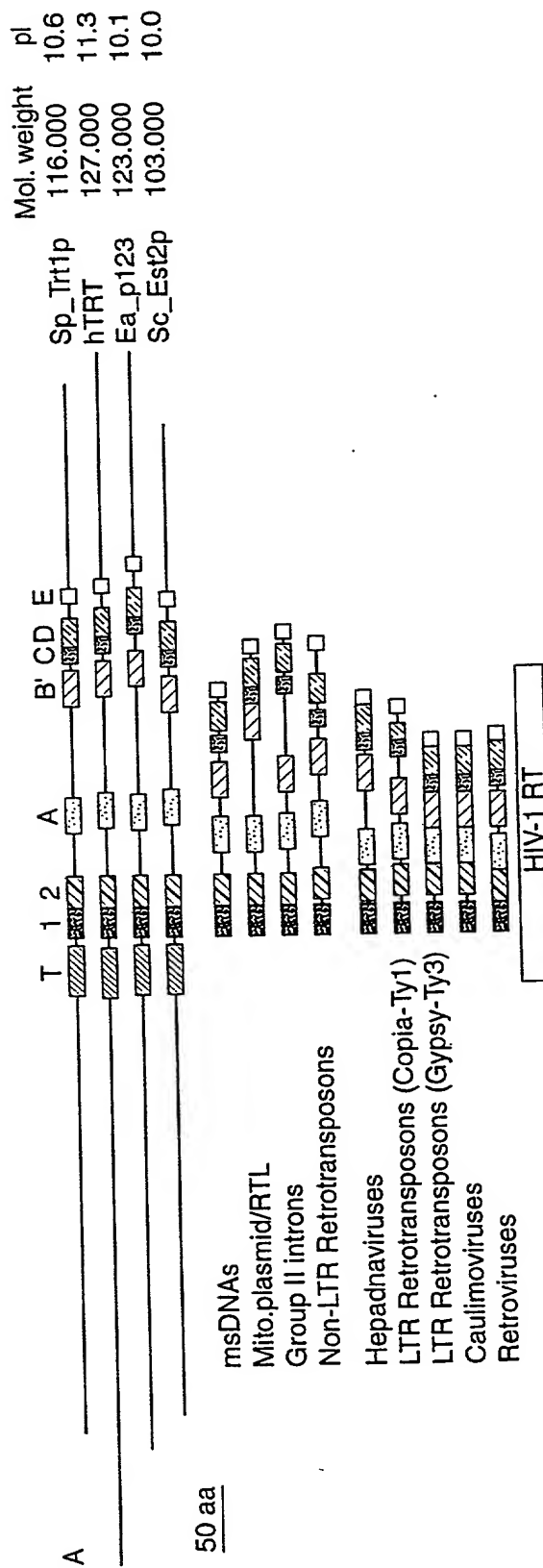


FIG. 2

3/34

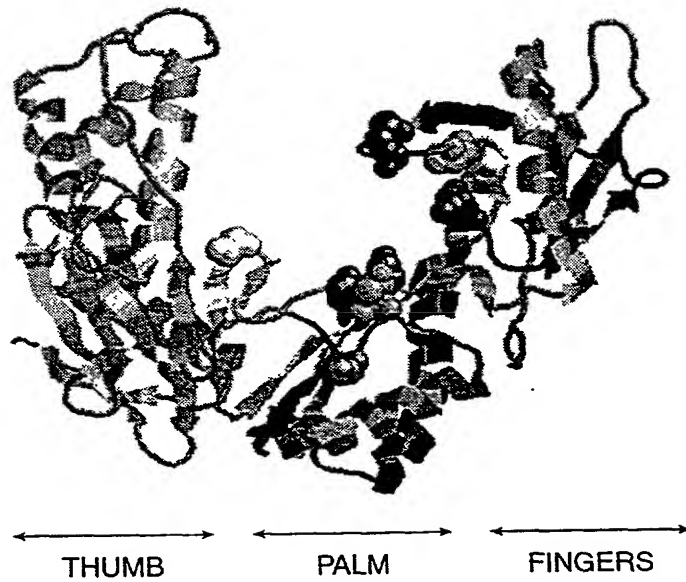


FIG. 3

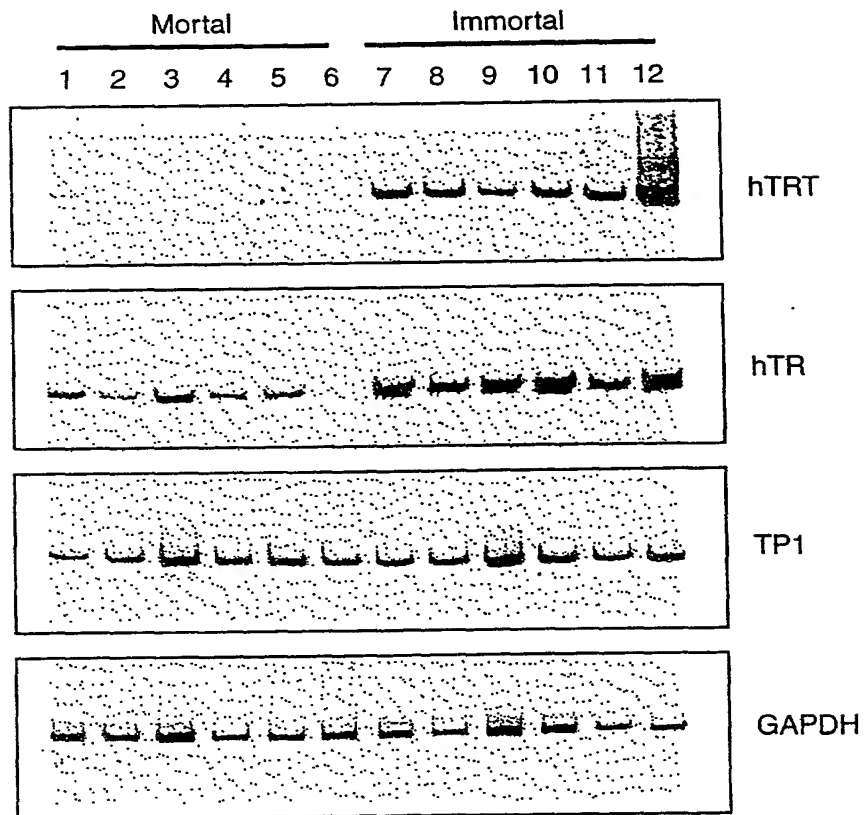


FIG. 5

FIG. 4

5/34

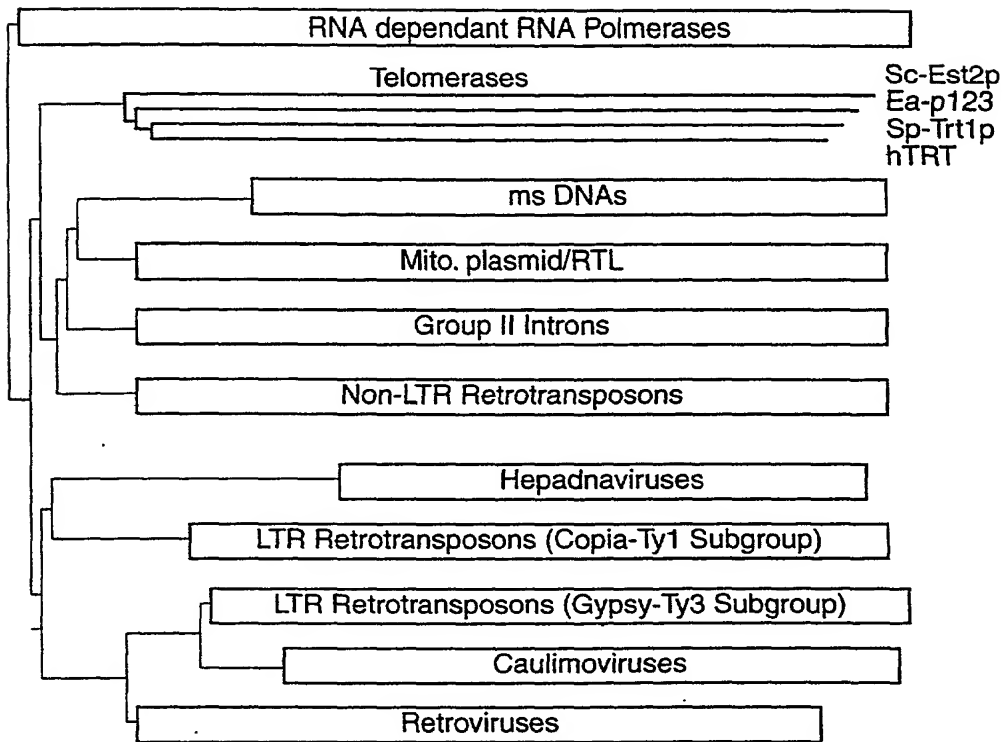


FIG. 6

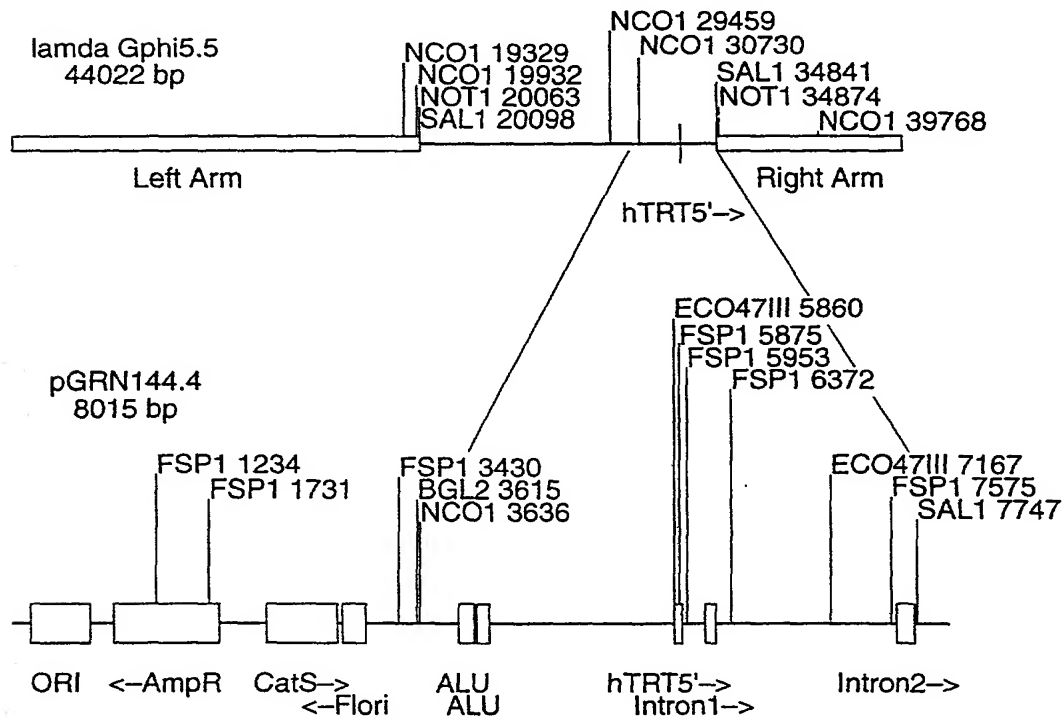


FIG. 7

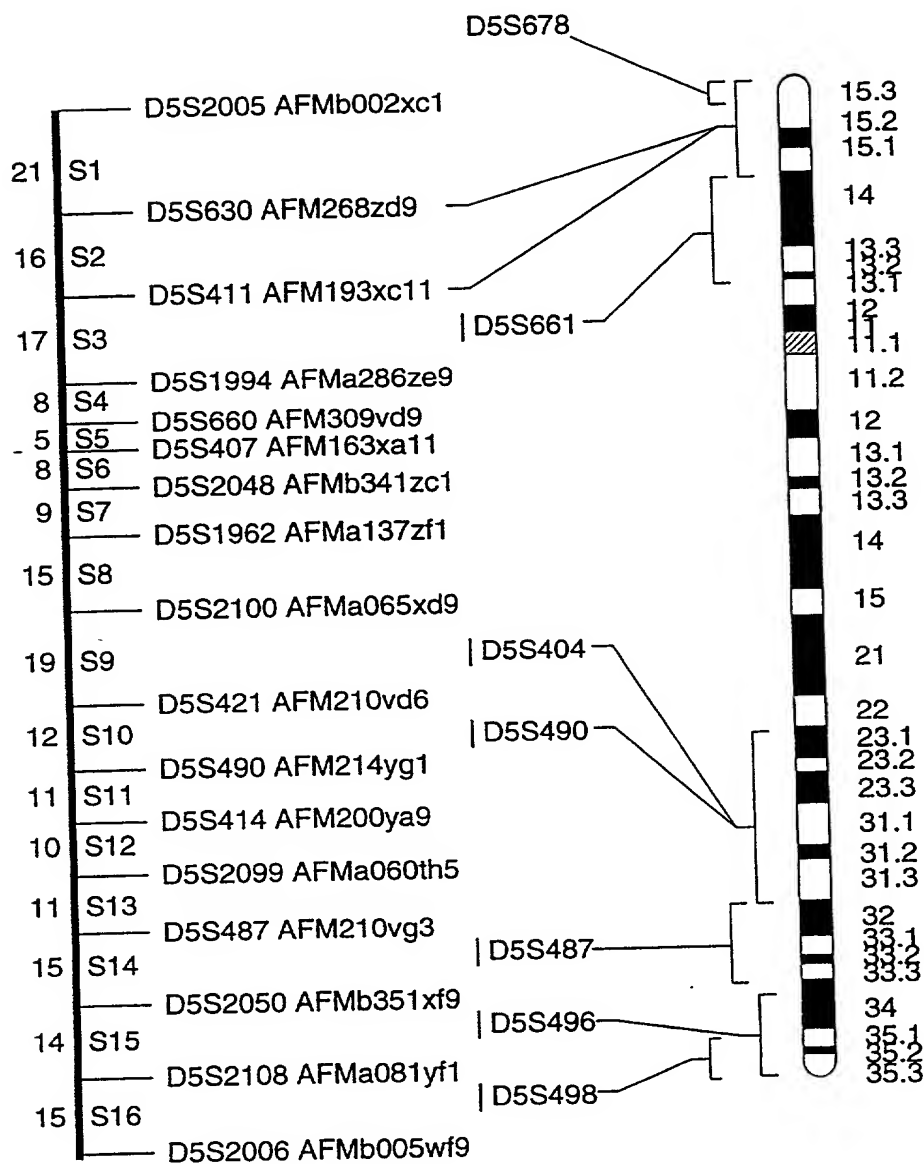


FIG. 8

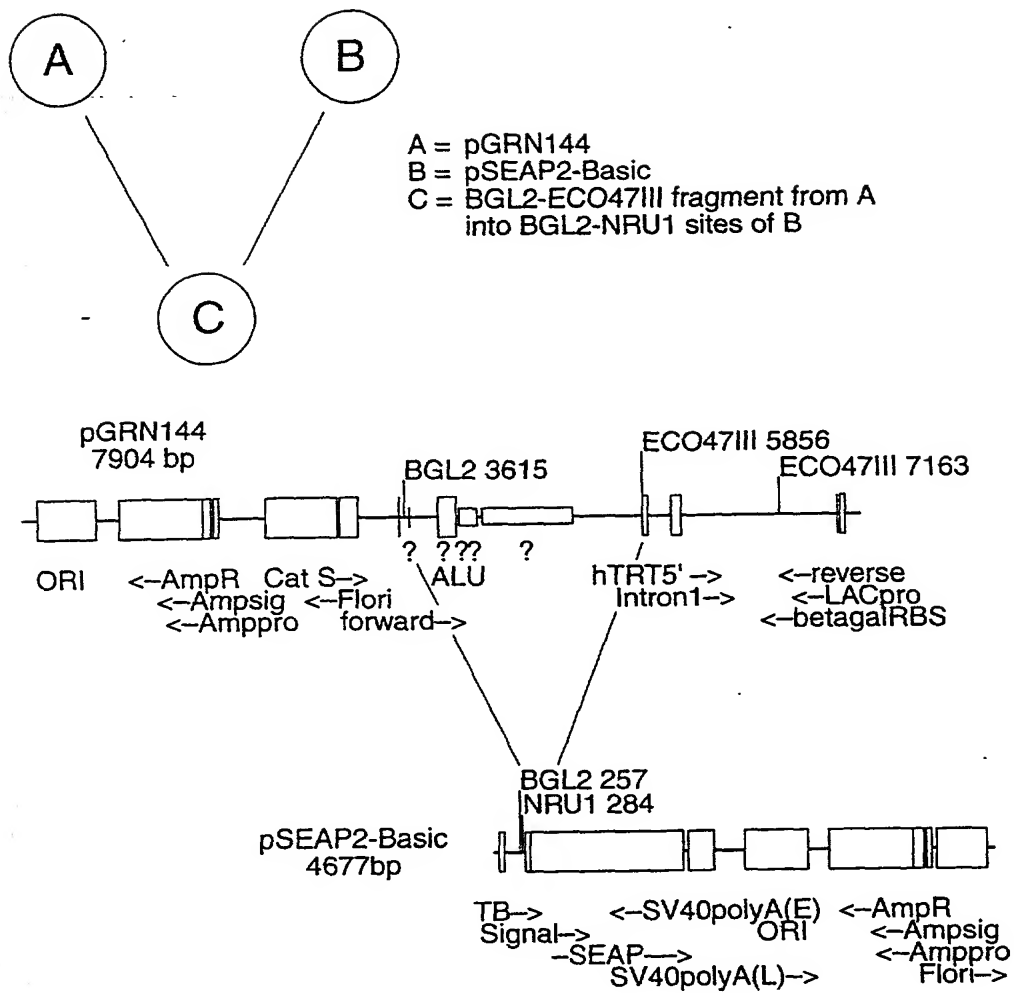


FIG. 9

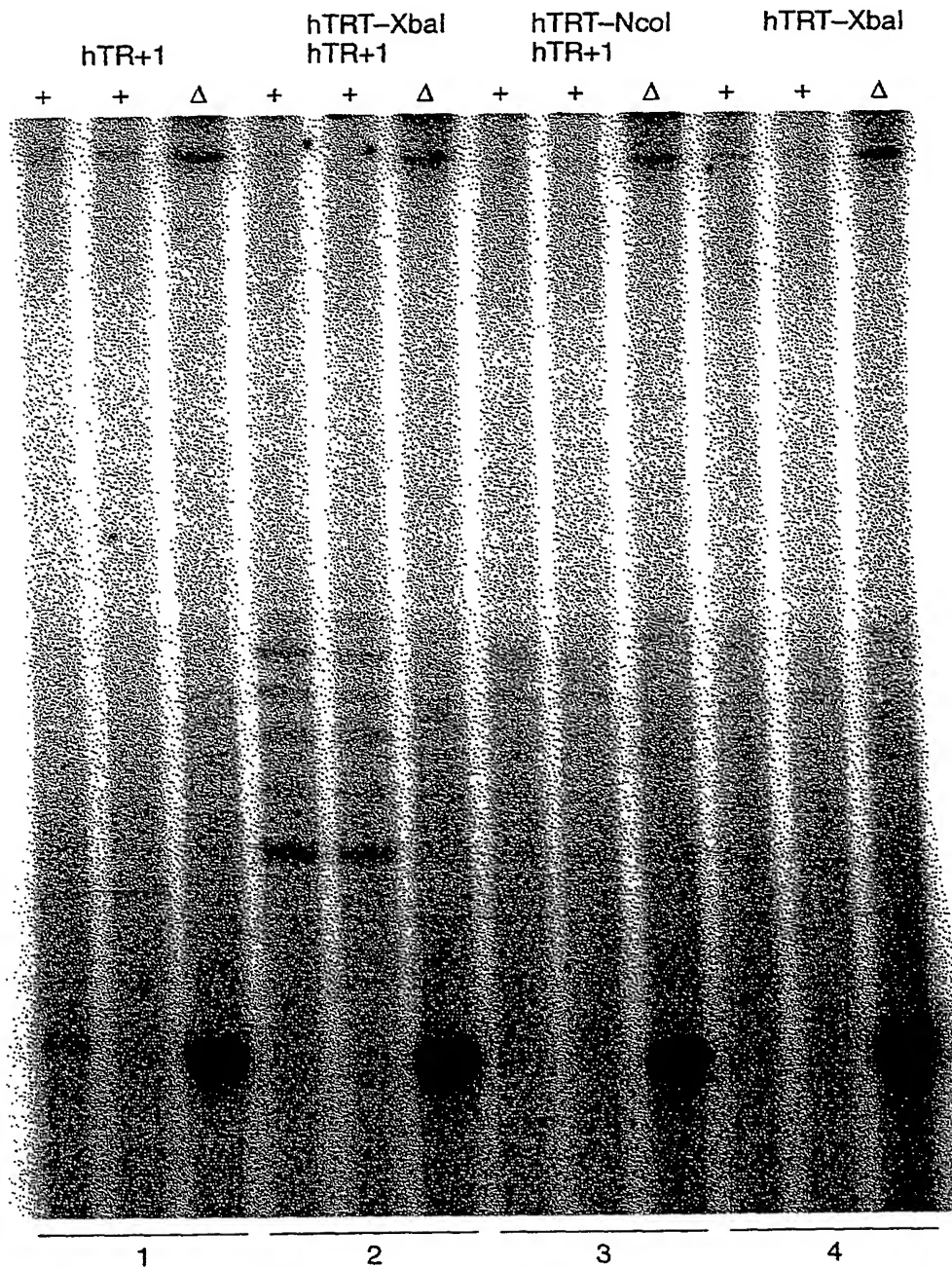


FIG. 10A

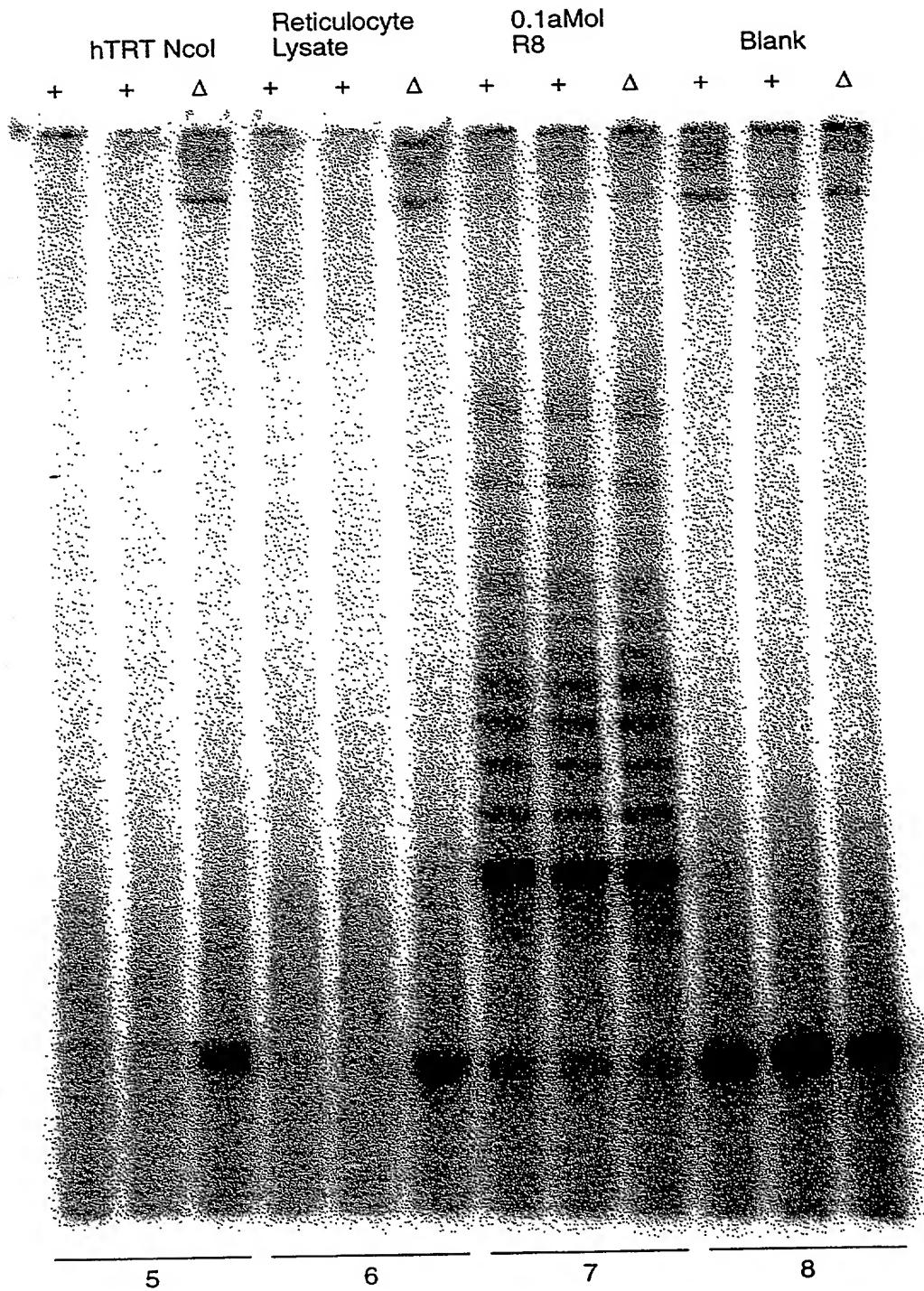


FIG. 10B

Telomerase Specific Motifs

	MOTIF T	MOTIF T'
TRT con	W1	E V
hTRT	FFY TE	I
spTRT	546 WLMSVVVVELLRSSFFVVTETTFQKNRLFYRKSVWSKLQSIGI	13 EAEVR
Ea_p123	429 WLYNSFIPIILQSFYITESDLNRNRTVYFRKDIWKLICRPF	12 ENNVR
Sc_Est2	441 WIFEDLVSLIRCFYVTEQQKSYKTYYYRKNIDVIMKMSI	12 EKEVE
	366 WLFRQLIPKIIQTFYCYCTEISSTVT.IVYFRHDTWNLITPFI	9 ENNVC

Telomerase RT Motifs (Fingers)

	MOTIF 1	MOTIF 2	MOTIF A	MOTIF B'
TRT con	R iPKk	fr I	p lyF D cyD i	Y q GiPQGs lS l Y
hTRT	11 SRLRFIPKPDG 0 LRPV	69 PELYFVKVDVTGAYDTI	104 YVQCQGIPOGSILSTLLCSLCY	
spTRT	10 AVIRLLPKKNT 0 FRLIT	66 RKKYFVRIDIKSCYDRI	99 YLQKVGIPQGSILSSFLCHFYM	
Ea_p123	10 GKLLIPKKT 0 FRPIM	67 PKLFFATMDIEKCYDSV	117 YKQTKGIPQGLCVSSILSSFY	
Sc_Est2	13 SKMRIIPKKS 2 FRIIA	68 PELYFMKFDVKSCYDSI	85 YIREDELFGQSSLSAPIVDLVY	
RT con	p hh h K	hR h	h hDh AF h	hpQG pP hh h
			GY	

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C	MOTIF D	MOTIF E
TRT con	l1lrl DdFL it	g n K	w g s l
hTRT	15 LLLRLVDDFLVIT	15 GVPEYGCVVNLKKTIV	24 WCGLLLDTRTL
spTRT	16 VLLRVVDDFLFIT	15 GFEKHNFSTLEKTVI	22 FFGFSVNMRS
Ea_p123	24 LLMRLTDDYLLIT	15 VSRENGFKFNKKLQ	28 WIGISIDMKTL
Sc_Est2	18 LILKLADDFLIIS	15 GFQKYNANRDKILA	25 WKHSSTMNPF
RT con	h Y DDhhh	Gh h cK h	hLG h
	F		

FIG. 11

181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
CCTGGGCGCGCGAAAGGCGCGGACACCGGGTCACGGACCACACGCACGGGACCCCTGCG

NFkB_CS1
GGGRQTYYYQC
NFkB-MHC-I.2
TGGGCTTCCCC

241 ACGGCCGCCCCCGCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCG
TGCCGGCGGGGGCGCGGGGGAGGAAGGCGGTCCACCCGGAGGGCCCCAGCCGCAGGC

Intron1

301 GCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
CGACCCCAACTCCCGCGGCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB_CS1
GGGRQTYYYQC
NFkB_CS2
RGGGRMTYYCC
Topo_II_cleavage_site
RNYNNCNGYNGKTNYNY

*****>

361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCCAGTGCTGCAGAGG
TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

```

1   AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
51  AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
251 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCa
451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
601 GCGAACTTCT GAAGGAACCTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTG TAATAGAAAT
851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT
901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TCGCTACATG
1001 CTTGAGAAAG TCAAAGATTT TAACCTCAAC TACTATTTAA CAAAATCTTG
1051 TCCTCTTCCA GAAAATTGGC GGGAAACGGA ACAAAAAATC GAAAACCTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
1151 TACACAACCTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
1251 AAAAGAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTAC
1301 AAAAATTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACATCT
1401 ACGTCTTATG GAAATGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451 CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTTGGATTC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACCAAAACCT CTTCTTTGCA
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
1951 AACATTCTTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAC
2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101 TGCATTGAA GGAGGACAAT ATCCAACCTT ATTCAAGTGT CTTGAAAATG
2151 AACAAAATGA CTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
2251 CCAATATAAT TACATTAAC TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT
2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

```

FIG. 13

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACTCT
2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901 AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC
2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001 CCTTGTTGTC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
3051 ACTTTTTCCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201 CTATTCTAAC TTATTTTGGG AAGTTAATTT TCAATTTTTG TCTTATATAC
3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

FIG. 13
(CONTINUED)

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGG
201 ADMNEPRCCS TCKYNVKNK DHFLNNINVP NWNMMKSRTN IFYCTHFNRN
251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIKKLK DKVIEKIAM
301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHელი
401 KNLLLEKINT REISWMQVET SAKHFYFFDH ENIYVLWKL RWIFEDLVVS
451 LIRCFYFVTE QQKSYSKTY YRKNIWDVIM KMSIADLKKE TLAEVQKEV
501 EEWKSLGFA PGKLRLIPKK TTFRPIMTFN KKIIVNSDRKT TKLTNTNKL
551 NSHMLMLTLK NRMFKDPFGF AVFNDDVMK KYEEFVCKWK QVGQPKLFFA
601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVDSKN
651 FRKKEMKDYF RQKFQKIALE GGOYPTLFSV LENEQNDLNA KKTLLIVEAKQ
701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFY
751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIKEL
801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
851 SIDMKTLALM PNINLRIEGL LCTNLNMQT KKASMWLKKK LKSFLMNNIT
901 HYFRKTITTE DFANKTLNKL FISGGYKMQ CAKEYKDHFK KNLAMSSMID
951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
1001 KYIFNRVCM I LKAKEAKLKS DQCQSLIQYD A

FIG. 14

1	ggtaacgatttacttcttcttcataagctaattgcttccctcgaacgctcctaataatctctggaaaataatttttacaaga	80
81	actcaataacaataccaagtcataattccaatatgaaggtgttattagtgatcgataataattcttattttatcggtcggtta	160
161	ccaagtataaggacaaaagaacaactcttcccctaagaacttttactttataatttacttttccaataataatttcg	240
241	ggtcgcgttacttttaactcgttggtactgttttagcttactcttagccaacccggcgtgtttctaccocgctcaatggatat	320
321	atgtccttgagtagtgcacagaaactccttaacaactctcttgatgagactatttagatccaatcacgctcggtcatattc	400
401	ttacatggagccttacacttttagatgagtcacgtcgcatgagtagtattgggtacatccaacgcttggtcgttgaagaag	480
481	gttgataattatttgcaaaatcatgtcctttagtgggtgaatccgcgaagaagtttttbgatgcttgcacacgctctagcatg	560
561	attgagataattcaaaaatttctatccactacaactcctttaacgggttttatttttctattttctatttcatgttgttt	640
641	ccaaatagtatcatctcgtatttaggctttttccggttttactcctggaaactgtaacctttttcactattcccccataatga	720
721	ataatctaaattagtttcgcgttataaattgtagtagtagtaagaagattgggtgattcactcgtgtgaatgttttagtttaa	800
801	gatacttttgcaaaactttattagctattcattatataaaaaaacctcctaataattataaataatttaacaaatttgcggtc	880
881	actattttataaaacggttatgatcagtaggacactttgcataatatagtattgcttaagtgttacttgaacttgc	958
959	ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CT'T CGC TTT CTA GAG AAT CAA TAT GTA	1018
1 M	T E H H T P K S R I L R F L E N Q Y V	20
1019	TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CT'T GTT TTG AGA GGG TCG CCG GCA AGC TCG	1078
21 Y	L C T L N D Y V Q L V L R G S P A S S	40
1079	TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT	1138
41 Y	S N I C E R L R S D V Q T S F S I F L	60
1139	CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA	1198
61 H	S T V V G F D S K P D E G V Q F S S P	80
1199	AAA TGC TCA CAG TCA GAG gtatatataattttgttttggatatttttctattcgggatatgctaataatattgggcag	1272
81 K	C S Q S E	86
1273	CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA	1332
87 L	I A N V V K Q M F D E S F E R R R N L	106
1333	CTG ATG AAA GGG TTT TCC ATG gtaaggattctctaattgtgaaataatttacctgcaattactgtttccaagaga	1405
107 L	M K G F S M	113
1406	ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT	1469
114	N H E D F R A M H V N G V Q N	128

FIG. 15

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
129 D L V S T F P N Y L I S I L E S K N W Q 148
1530 CTT TTG TTA GAA AT gtaaataccggttaagatgttgcgactttgaaacagactgacaagtatag T ATC GGC 1601
149 L L L E I 155
1602 AGT GAT GCC ATG CAT TAC TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
156 S D A M H Y L L S K G S I F E A L P N D 175
1662 AAT TAC CTT CAG AAT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
176 N Y L Q I S G I P L F K N N V F E E T V 195
1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
196 S K K R K R T I E T S I T Q N K S A R K 215
1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
216 E V S W N S I S I S R F S I F Y R S S Y 235
1842 AAG AAG TTT AAG CAA G gtaactaatactgttattccttcataactaatttag AT CTA TAT TTT AAC 1907
236 K K F K Q D L Y F N 245
1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
246 L H S I C D R N T V H M W L Q W I F P R 265
1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
266 Q F G L I N A F Q V K Q L H K V I P L V 285
2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
286 S Q S T V V P K R L L K V Y P L I E Q T 305
2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
306 A K R L H R I S L S K V Y N H Y C P Y I 325
2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
326 D T H D D E K I L S Y S L K K P N Q V F A 345
2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
346 F L R S I L V R V F P K L L I W G N Q R I 365

FIG. 15
(CONTINUED)

2268 TTT GAG ATA ATA TTA AAA G gatttgataaaatttattaccactaacgattttaccag AC CTC GAA ACT 2336
366 F E I I L K D L E T 375
2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
376 F L K L S R Y E S F S L H Y L M S N I K 395
2397 gtaatatgccaaattttttaccatttaataacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
396 I S E I E W L V L G 405
2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
406 K R S N A K M C L S D F E K R K Q I F A 425
2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
426 E F I Y W L Y N S F I I P I L Q S F F Y 445
2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
446 I T E S S D L R N R T V Y F R K D I W K 465
2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
466 L L C R P F I T S M K M E A F E K I N E 485
2706 gatttttaaaagtatttttgcaaaagctaatttttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
486 N N V R M D T Q K T 495
2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
496 T L P P A V I I R L L P K K N T F R L I T 515
2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaatttttggtcatcaatgtactttacttctaatttata 2906
516 N L R K R F L I K 524
2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
525 M G S N K K M L V S T N Q T L R P V 542
2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
543 A S I L K H L I N E E S S G I P F N L E 562
3028 GTT TAC ATG AAG CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 15
(CONTINUED)

17/34

3089 tatataatgcgcgattcctcattatttaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
 582 R K K Y F V R I D I 591
 3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
 592 K S C Y D R I K Q D L M F R I V K K L 611
 3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
 612 K D P E F V I R K Y A T I H A T S D R A 631
 3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagttttattttttcattggaattttttaacaa 3343
 632 T K N F V S E A F S Y F 643
 3344 attcttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
 644 D M V P F E K V Q L L S M K T 659
 3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
 660 S D T L F V D F V D Y W T K S S E I F 679
 3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgtgaattgtaataaca 3532
 680 K M L K E H L S G H I V K 692
 3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593
 693 I G N S Q Y L Q K V G I P Q G S 708
 3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
 709 I L S S F L C H F Y M E D L I D E Y L S 728
 3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
 729 F T K K K G S V L L R V V D D F L F I T 748
 3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtaggttgctgtcattcc 3777
 749 V N K K D A K K F L N L S L R G 764
 3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
 765 F E K H N F S T S L E K T V 778
 3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900
 779 I N F E N S N G I I N N T F F N E S K K 798

FIG. 15
(CONTINUED)

CUF 30 " SEQUENCE

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
799 R M P F F G F S V N M R S L D T L L A C 818
3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
819 P K I D E A L F N S T S V E L T K H M G 838
4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataatagctgacaaataatcag A TCG 4089
839 K S F F Y K I L R S 848
4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
849 S L A S F A Q V F I D I T H N S K F N S 868
4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
869 C C N I Y R L G Y S M C M R A Q A Y L K 888
4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
889 R M K D I F I P Q R M F I T D 903
4275 aaagtcattaattaaccccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
904 L L N V I G R K I W K L A 917
4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
918 E I L G Y T S R F L S S A E V K W 935
4402 ggtctcgagacttcagcaatatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
936 L F C L G M R D G L K 946
4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
947 P S F K Y H P C F E Q L I Y Q F Q S L T 966
4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA ATA 4588
967 D L I K P L R P V L R Q V L F L H R R I 986
4589 GCT GAT TAA tgtcatttttcaatttattattatatacatccttttattactgggtgtctttaaacaaatattattactaagtata 4665
987 A D * 989

FIG. 15
(CONTINUED)

FIG. 15
(CONTINUED)

1 gcagcgctgc gtctgtctgc gcacgtggga agccctggcc ccggccaccc ccgcgatgcc
 61 gcgcgctccc cgctgccgag ccgtgcgctc cctgctgcgc agccactacc gcgaggtgct
 121 gccgctggcc acgttcgtgc ggcgcctggg gccccagggc tggcggtgg tgcagcgcg
 181 ggacccggcg gctttccgcg cgctgggtgg ccagtgcctg gtgtgcgtgc cctgggacgc
 241 acggccgccc ccgcgcgccc cctccttccg ccaggtgtcc tgctgaagg agctggtggc
 301 ccgagtgctg cagaggctgt gcgagcgcg cgcaagaac gtgctggcct tccgcttcgc
 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta
 421 cctgcccac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgcgtgctgc
 481 ccgcgtgggc gacagctgc tggttcacct gctggcacgc tgcgcgtct tttgtctggt
 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgcgctg taccagctcg gcgctgccac
 601 tcaggcccgg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc
 661 ctggaaccat agcgtcaggg aggcgggggt cccctgggc ctgccagccc cgggtgcgag
 721 gaggcgcggg ggcagtgcca gccgaagtct gccgttgccc aagaggccca ggcgtggcgc
 781 tgcccctgag ccggagcgga cgcccgttgg gcaggggtcc tggggccacc cgggcaggac
 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccg ccgaagaagc
 901 cacctctttg gaggtgctgc tctctggcac gcgccactcc caccatccg tgggcgccca
 961 gcaccacgcg gggcccccat ccacactgcg gccaccact ccctgggaca cgccttgtcc
 1021 cccggtgtac gccgagacca agcacttctt ctactcctca ggcgacaagg agcagctgcg
 1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga
 1141 gaccatcttt ctgggttcca ggcctctgat gccagggact ccccgagggt tgcctgctga
 1201 gccccagcgc tactggcaaa tgccgcccct gtttctggag ctgcttggga accacgcgca
 1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgtg cgagctgcgg tcaccccagc
 1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga
 1381 cacagacccc cgctgcctgg tgcagctgct ccgccagcac agcagccctt ggcagctgta
 1441 cggcttctgt cgggcctgccc tgccgcccgt ggtgccccca ggcctctggg gctccaggca
 1501 caacgaacgc cgcttctca ggaacaccaa gaagtctatc tccctgggga agcatgccaa
 1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgag
 1621 gagcccaggg gttggctgtg ttccggccgc agagcacctg ctgctgagg agatcctggc
 1681 caagtctctg tactgtctga ctactgtgta cgtcgtcgag ctgctcaggt ctttctttta
 1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag
 1801 caagtgtcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggagctgtc
 1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg
 1921 cttcatcccc aagcctgacg ggcctggccc gattgtgaac atggactacg tcgtgggagc
 1981 cagaacgttc cgcagagaaa agaggggcga gcgtctcacc tcgagggtga aggcactgtt
 2041 cagcgtgctc aactacgagc gggcgcgggc ccccggcctc ctggggcgct ctgtgtggg
 2101 cctggacgat atccacaggg cctggcgcac ctctgtgctg cgtgtgcggg ccaggagacc
 2161 gccgcctgag ctgtactttg tcaagggtga tgtgacgggc gcgtacgaca ccatccccc
 2221 ggacaggctc acggagggtc tccagcagc catcaaacc cagaacacgt actgctgctg
 2281 tcggtatgcc gtggtccaga aggcgcgcca tgggcacgtc cgcaaggcct tcaagagcca
 2341 cgtctctacc ttgacagacc tccagccgta catcgacag ttcgtggctc atgaggccag
 2401 gaccagcccg ctgagggatg ccgtgcctat cgagcagagc tcctccctga atgaggccag
 2461 cagtggcctc ttcgacgtct tcctacgctt catgtgccac caccgctgct gcactcaggg
 2521 caagtccctac gtccagtgc aggggatccc gcagggtctc atcctctcca cgtgctctg
 2581 cagcctgtgc tacggcgaca tggagaacaa cgtgtttgcy gggattccgg cagcgaaaac
 2641 gctcctgctg ttggtggatg atttcttgtt ggtgacacct cacctcacc acgcgaaac
 2701 cttcctcagg accctggctc gaggtgtccc tgagtatggc tgcgtgggtg acttgcgga
 2761 gacagtgggt aacttccctg tagaagacga ggcctgggtt ggcacggctt ttgttcagat
 2821 gccggcccac ggcctatttc cctgtgctg cctgctgctg gatacccggg ccttgagggt
 2881 cgagagcgac tactccagct atgcccggac ctccatcaga gccagtctca ccttcaaccg
 2941 cggcttcaag gctgggagga acatgcgtcg caaactcttt ggggtcttgc ggctgaagtg
 3001 tcacagcctg tttctggatt tgcaggtgaa cagcctccag acggtgtgca ccaacatcta
 3061 caagatcctc ctgctgcagg cgtacaggtt tcacgcatgt gtgctgcagc tcccatttca
 3121 tcagcaagtt tggagaacc ccacattttt cctgcgcgtc atctctgaca cggcctccct
 3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctggggggcca agggcgccgc
 3241 cggccctctg ccctccgagg ccgtgcagt gctgtgccac caagcattcc tgctcaagct
 3301 gactgcacac cgtgtcacct acgtgcacgt cctggggtca ctacggacag cccagacgca
 3361 gctgagtcgg aagctcccgg ggacgacgt gactgccctg gaggccgag ccaacccggc
 3421 actgccctca gacttcaaga ccatcctgga ctgatggcca cccggccaca gccaggccga
 3481 gagcagacac cagcagccct gtcacgcgg gctctacgtc ccagggaggg agggcgggcc
 3541 cacacccagg ccgcacgcgc tgggagctct aggcctgagt gagtgtttgg ccgaggcctg
 3601 catgtccggc tgaaggctga gtgtccggct gaggcctgag cgagtgtcca gccaaaggct
 3661 gagtgtccag cacacctgcc gtcttcaact ccccacaggc tggcgctcgg ctccaccca
 3721 gggccagctt ttcttcacca ggagccgggc ttccactccc gccttccacc tagtccatc
 3781 ccagattcgc cattgttcac cctgcgccct gccctccttt gccttccacc cccaccatcc
 3841 aggtggagac cctgagaagg accctgggag ctctgggaat ttggagtgac caaagggtg
 3901 ccctgtacac aggcgaggac cctgcacct gatgggggtc cctgtgggtc aaattggggg
 3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa

FIG. 16

21/34

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLLGPQGWRLVQRGDP
AAFRALVAQCLVCPWDARPPPAAPSFRQVSCLKELVARVLQRL
CERGAKNVLAFGFALLDGARGGPPPEAFTTSVRSYLPNTVTDALR
GSGAWGLLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY
QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG
ARRRGGASASRLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRG
PSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP
STSRRPPRPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLLSSLRP
SLTGARRLVETIFLGSRPWMPGTFRRLPRLPQRYWQMRPLFLEL
LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPPQSVAAPPEE
EDTDPRRLVQLLRQHSSPWQVYGFVRACLRLVPPGLWGSRHNE
RRLFLRNTKFIISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC
VPAEHLRLREEILAKFLHWLMSVYVVELLRSFFVYTETTFQKNR
LFFYRKSWSKLQSIGIRQLKRVQLRELSAEVQRHREARPAL
LTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKA
LFSVLNHERARRPGLLGASVLGLDDIHRWRTFVLRVRAQDPPP
ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ
KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPOGSI
LSTLLCSLCYGD MENKLFAGIRRDGLLLRLVDDFLLVTPHLTHA
KTFLRLTLVRGVPEYGCVVNLRLKTVVNF PVEDEALGGTAFVQMPA
HGLFPWCGLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
NMRRLKFGVRLRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRF
HACVLQLPFHQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL
GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGLSLRTAQ
TQLSRKLPGTTLTALEAAANPALPSDFKTILD

FIG. 17

GGCCAAGTTCTCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTT
TTATGTACAGGAGACCAGCTTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTG
GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT
GTCGGAAGCAGAGGTGAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACT
CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCGGATTGTGAACATGGACTACGTCGTGGG
AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCAT
GTTGAGCGTGTCTCAACTCAGAGCGGGCGCGGCGCCCCGGCCTCTGGGCGCCTCTGTGCT
GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGA
CCCGCCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
CCAGGACAGGCTCAGCGAGGTGATCGCCAGCATCATCAAACCCAGAACACGTACTGCGT
GCGTCGGTATGCGTGGTCCAGAAAGCCGCCCCATGGGCACGTCCGCAAGGCCTTCAAGAG
CCAGCTCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT
GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTTCGCGGGGATTTCGGCGGGACGGG
TGCTCCTGCGTTCGTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCACGCGAAAA
CCTTCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACCTGCGGA
AGACAGTGGTGAACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGGCTTTTGTTCAGA
TGCCGGCCACGGCCTATTCCCCTGGTGGGCTGCTGCTGGATACCCGGACCCTGGAGG
TGCAGAGCGACTACTCCAGCTATGCCCCGACCTCCATCAGAGCCAGTCTCACCTTCAACC
GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTTCGGCTGAAGT
GTCACAGCCTGTTTCTGGATTTGCAAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT
ACAAGATCCTCTCTGCTGCAAGGCTACAGGTTTTCACGCATGTGTGCTGCAGCTCCCATTTT
ATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGCGTCATCTCTGACACGGCCTCCC
TCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTGCTGCGGGGCCAAGGGCGCCG
CCGGCC7TCTGCCCTCCGAGGCCGTGCACTGGCTGTGCCACCAAGCATTCCTGCTCAAGC
TGACTCGACACCGTGTACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGC
AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGAGCCAACCCGG
CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCG
AGAGCAGACACCAGCAGCCCTGTACGCGCGGGCTCTACGTCCAGGGAGGGAGGGGCGGC
CCACACCCAGGCTGACCGCTGGGAGTCTGAGGCTGAGTGAGTGTTCGCGGAGGCGCT
GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGCGAGTGTCCAGCCAAGGGC
TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
AGGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC
CCCAGATTCCGCAATTGTTACCCCCCGCCCTGCCCTCCTTTGCCCTTCCACCCCCACCATC
CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAGAGGTGT
GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTTGGGG
GGAGGTGCTGTGGGAGTAAATACTGAATATATGAGTTTTTTCAGTTTTTG0AAAAAAAAA
AAAAAAAAAAAAAAAAA

FIG. 18

22/34

MetSerValTyrValValGluLeuLeuArgSerPhePhe
TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe
PheTyrArgLysSerValTrpSerLysLeuGlnSerIle
GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu
LeuSerGluAlaGluValArgGlnHisArgGluAlaArg
ProAlaLeuLeuThrSerArgLeuArgPheIleProLys
ProAspGlyLeuArgProIleValAsnMetAspTyrVal
ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu
ArgLeuThrSerArgValLysAlaLeuPheSerValLeu
AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla
SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg
ThrPheValLeuArgValArgAlaGlnAspProProPro
GluLeuTyrPheValLysValAspValThrGlyAlaTyr
AspThrIleProGlnAspArgLeuThrGluValIleAla
SerIleIleLysProGlnAsnThrTyrCysValArgArg
TyrAlaValValGlnLysAlaAlaHisGlyHisValArg
LysAlaPheLysSerHisValLeuArgProValProGly
AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln
ProValLeuArgArgHisGlyGluGlnAlaValCysGly
AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

FIG. 19

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGCCACCCCCGCG met
ATG
10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC
20 30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG
40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT
50 60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC
70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC
80 90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC
100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG
110 120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

FIG. 20

23/34

130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

140
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

150
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

160
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

170
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

180
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

190
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

200
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

210
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

220
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

230
val val ser pro ala arg pro ala glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

240
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

250
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

260
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

270
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

FIG. 20
(CONTINUED)

24/34

340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

360
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

370
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

380
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

390
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

400
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

410
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

420
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

430
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

440
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

450
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

460
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

470
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

480
530
540

FIG. 20
(CONTINUED)

550

560

570

580

590

610

620

640

650

670

680

700

710

730

740

750

FIG. 20

26/34

760
his gly his val arg lys ala phe lys ser his val leu arg pro
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

770 780
val pro gly asp pro ala gly leu his pro leu his ala ala leu
GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

790
gln pro val leu arg arg his gly glu gln ala val cys gly asp
CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

800 807
ser ala gly arg ala ala pro ala phe gly gly OP
TCG GGG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTGTTGGT

GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA
GTATGGCTGCGTGGTGAACCTTGCGGAAGACAGTGGTGAACCTTCCTGTAGAAGACGAGGC
CCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCACGGCCTATTCCCCTGGTGCGGCCT
GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC
CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA
ACTCTTTGGGGTCTTGCGGCTGAAGTGTACAGCCTGTTTCTGGATTTCAGGTGAACAG
CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA
CGCATGTGTGCTGCAGCTCCCATTTTCATCAGCAAGTTTGGAAGAACCCACATTTTTCCT
GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG
GATGTGCTGCGTGGGGGCCAAGGGCGCCGCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCT
GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCCT
GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC
TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG
ATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTCACGCCGGGCT
CTACGTCCCAGGGAGGGAGGGGCGGCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG
CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG
GCCTGAGCGAGTGTCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC
CACAGGCTGGCGCTCGGCTCCACCCAGGGCCAGCTTTTTCCTCACCAGGAGCCCGGCTTC
CACTCCCCACATAGGAATAGTCCATCCCCAGATTGCGCCATTGTTACCCCTCGCCCTGCC
CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC
TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT
GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAATACTGAATATATG
AGTTTTTTCAGTTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 20
(CONTINUED)

FIG. 21

28/34

4321 GGTGTTTTTAAGCCAATNANAAAAATTTTNTATGTTGTTNNNNNNNNNNNNNNNNNNNN
CCACAAAAATTCGGTTANTNTTTTAAAAAANTACAACAAANNNNNNNNNNNNNNNNNNNNN

4381 NNN
NN

4441 NNN
NN

4501 NNN
NN

4561 NNN
NN

4621 NNN
NN

4681 NNN
NN

4741 NNN
NN

4801 NNN
NN

4861 NNN
NN

4921 NNN
NN

4981 NNN
NN

5041 NGCCANGRAGGGGGCCAGGTTCCAANTTCCCAACCKTTTTWGGARGGACNGCCCCCAGGG
NCGGTNCYTCCCCCGGTCCAAGGTTNAAGGGTTGMAAAAWCCTYCCTGNCGGGGGTCCC

5101 GGGGATRAACAGANTNGGGGKGGTWGGGTTNAKGGTGGGAACNCCTTNGCGCCTGGAG
CCCCTAYTTGTCTNANCCCCMCCAWCCCAANTMCCACCCCTGNGGAANCGSCGGACCTC

5161 AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG
TTGCACGTTTCTCCTTTACTTCCCGGACMCAGTTCCCTCGGGTTCANCCGCCCCYTCAAAC

5221 CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCTTCGGG
GTCCCTCCGTGAGGCCCTCCAGGSCGCACGGGCAGGTTCCTCGTTACGCAGGAAGCCC

5281 TTCGTCCCCAWGCCGCGTCTACGCGCTYCCGTCTCCCTTCACGTTCCGGCATTCGTG
AAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC

5341 GTGCCCGGAGCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGATCAG
CACGGGCCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTGCGGACCCAGAGGCCTAGTC

5401 GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCAGGGCCTCCACATCATGGCCCCCT
CGGTCGCCGGTTTCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCGGGGA

FIG. 21
(CONTINUED)

29/34

5461 CCCTCGGGTTACCCACAGCCTAGGCCGGATTTCGACCTCTCTCCGCTGGGGCCCTCGCCT
GGGAGCCCAATGGGGTGTTCGGATCCGGCCTAAGCTGGAGAGAGGCGACCCCGGGAGCGGA

Sp1

5521 GGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGGGCGGGGAAGCGCGGCCCATACCC
CCGCAGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCTTCGCGCCGGGTATGGG

5581 CCGGGTCCGCCCCGAAGCAGCTGCGCTGTTCGGGGCCAGGCCGGGCTCCCAGTGGATTTCGC
GGCCAGGCGGGCCTTCGTTCGACGCGACAGCCCGGTCCGGCCCGAGGGTCACCTAAGCG

Topo_II_cleavage_site

5641 GGGCACAGACGCCCAGGACCGCGCTTCCACGTGGCGGAAGGACTGGGGACCCGGGCACC
CCCGTGTCTGCGGGTCTCTGGCGCGAAGGGTGCACCGCCTTCCTGACCCCTGGGCCCCTGG

E2F

5701 CGTCCTGCCCCCTTCACCTTCCAGCTCCGCTTCTTCCGCGCGGACCCGGCCCCGTCCCGAA
GCAGGACGGGGAAGTGAAGGTTCGAGGCGAAGAAGGCGCGCCTGGGCGGGGCGAGGGCTT

E

5761 CCCTTCCCAGGTCCCGGCCAGCCCTTCCGGGCCCTCCAGCCCTCCCCTTCTTTTC
GGGAAGGGTCCAGGGCCGGGTTCGGGAAGGCCCGGGAGGGTTCGGGAGGGGAAGGAAAG

Sp1
=====

2F		NFkB		h
*****		*****	*****	
5821	CGCGGCCCGCCCTCTCCTTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCTGCTGCGCA			
	GCGCCGGGGCGGGAGAGGAAGCGCCGCGCTCAAAGTCCGTCGCGACGCAGGACGACGCGT			
		^	^	
		5860	5875	
		ECO47III	FSP1	

TRT5'
*****>

5881 CGTGGGAAGCCCTGGCCCCGGCCACCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCG
GCACCCTTCGGGACCGGGGCCGGTGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTCGGC

5941 TGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTTCGTGCGGC
ACGCGAGGGACGACGCGTCCGGTGATGGCGCTCCACGACGGCGACCGGTGCAAGCACGCGC

^
5953
FSP1

6001 GCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGACCCGGCGGCTTTCCGCGCGC
CGGACCCCGGGTCCCACCGCCGACCACGTTCGCGCCCTGGGCCGCCGAAAGGCGCGCG

6061 TGGTGGCCCCAGTGCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCGCGCCCCCT
ACCACCGGGTCACGGACCACACGCACGGGACCCTGCGTGCCGGCGGGGGCGGGGGGA

NFkB
=====

FIG. 21
(CONTINUED)

30/34

6121 CCTTCCGCCAGGTGGGCTCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGG
GGAAGGCGGTCCACCCGAGGGGCCCCAGCCGAGGCCGACCCCAACTCCCGCCGGCCCC

Topo_II_cleavage_s

:::::::::::::::

NFkB

=====

Intron1

*****>

6181 GGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGTGTC
CCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAGTCCCGCGAAGGGGGCGTCCACAG

ite

:

6241 CTGCCTGAAGGAGCTGGTGGCCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGCGAAGAA
GACGGACTTCCTCGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTTCTT

6301 CGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCCGCGGGGGCCCCCGAGGCCTT
GCACGACCCGAAGCCGAAGCGCGACGACCTGCCCCGGGCGCCCCCGGGGGGGCTCCGGAA

6361 CACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGG
GTGGTGGTCGCACGCGTCGATGGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTCGCC

^

6372

FSP1

6421 GGCGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTTACCTGCTGGCAGC
CCGACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCGTGC

6481 CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCGCT
GACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCCGGCGCGA

6541 GTACCAGCTCGGCGCTGCCACTCAGGCCCCGCCCCCGCCACACGCTAGTGGAACCCGAAG
CATGGTCGAGCCGCGACGGTGAGTCCGGGCGGGGGCGGTGTGCGATCACCTGGGGCTTC

6601 GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGG
CGCAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCC

6661 CCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCC
GGACGGTCGGGGCCCCACGCTCCTCCGCGCCCCCGTACGGTCGGCTTCAGACGGCAACGG

6721 CAAGAGGCCCCAGGCGTGGCGCTGCCCCCTGAGCCGGAGCGGACGCCCCGTTGGGCAGGGGTC
GTTCTCCGGGTCCGCACCGCGACGGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAG

6781 CTGGGCCCCACCGGGCAGGACGCGTGACCGAGTGACCGTGGTTTCTGTGTGGTGTACCC
GACCCGGGTGGGCCCCGTCTTGCACCTGGCTCACTGGCACCAAGACACACCACAGTGG

6841 TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC
ACGGTCTGGGCGGCTTCTTCGGTGGAGAACTCCACGCGAGAGACCGTGCGCGGTGAG

6901 CCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCATCCACATCGCGGCCACCACG
GGTGGGTAGGCACCCGGCGGTGCTGGTGCGCCCGGGGGGTAGGTGTAGCGCCGCTGGTGC

FIG. 21
(CONTINUED)

31/34

6961 TCCCTGGGACACGCCTTGTCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC
AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAG
7021 AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCTACTCAGCTCTCTGAGGCCAGCCTGAC
TCCGCTGTTCTCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTTCGACTG
7081 TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGGAC
ACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTG
7141 TCCCCGCAGGTTGCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCTGTTTCTGGA
AGGGGCGTCCAACGGGGCGGACGGGGTTCGCGATGACCGTTTACGCCGGGGACAAAGACCT

7167
EC047III

7201 GCTGCTTGGGAACCACGCGCAGTGCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCT
CGACGAACCCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGCGA
7261 GCGAGCTGCGGTACCCCCAGCAGCCGGTGTCTGTGCCCCGGGAGAAGCCCCAGGGCTCTGT
CGCTCGACGCCAGTGGGGTCGTGCGCCACAGACACGGGCCCTCTTCGGGGTCCCGAGACA
7321 GGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCAGCA
CCGCCGGGGGCTCCTCCTCTGTGTCTGGGGGACGCGGACCACGTCGACGAGGCGGTCTGT
7381 CAGCAGCCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGTCCCC
GTCGTGCGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGGGGG
7441 AGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCAT
TCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTGGTTCTTCAAGTA
7501 CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCG
GAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCACGC

7561 GGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTGAGGGGCCAGG
CCTGACGCGAACCAGCGCTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGGTCC

7575
FSP1

Intron2

7621 CCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCTCCT
GGGGTCTCGACTTACGTCATCCCCGAGTCTTTTCCCCCGTCCGTCTCGGGACCAGGAGGA

7681 GTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTGAGTGGACACGGTG
CAGAGGTAGCAGTGACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC

**>

7741 ATCGAGGTGCACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATA
TAGCTCCAGCTGAGATCTCCTAGGGGGCCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT

7747
SAL1

FIG. 21
(CONTINUED)

32/34

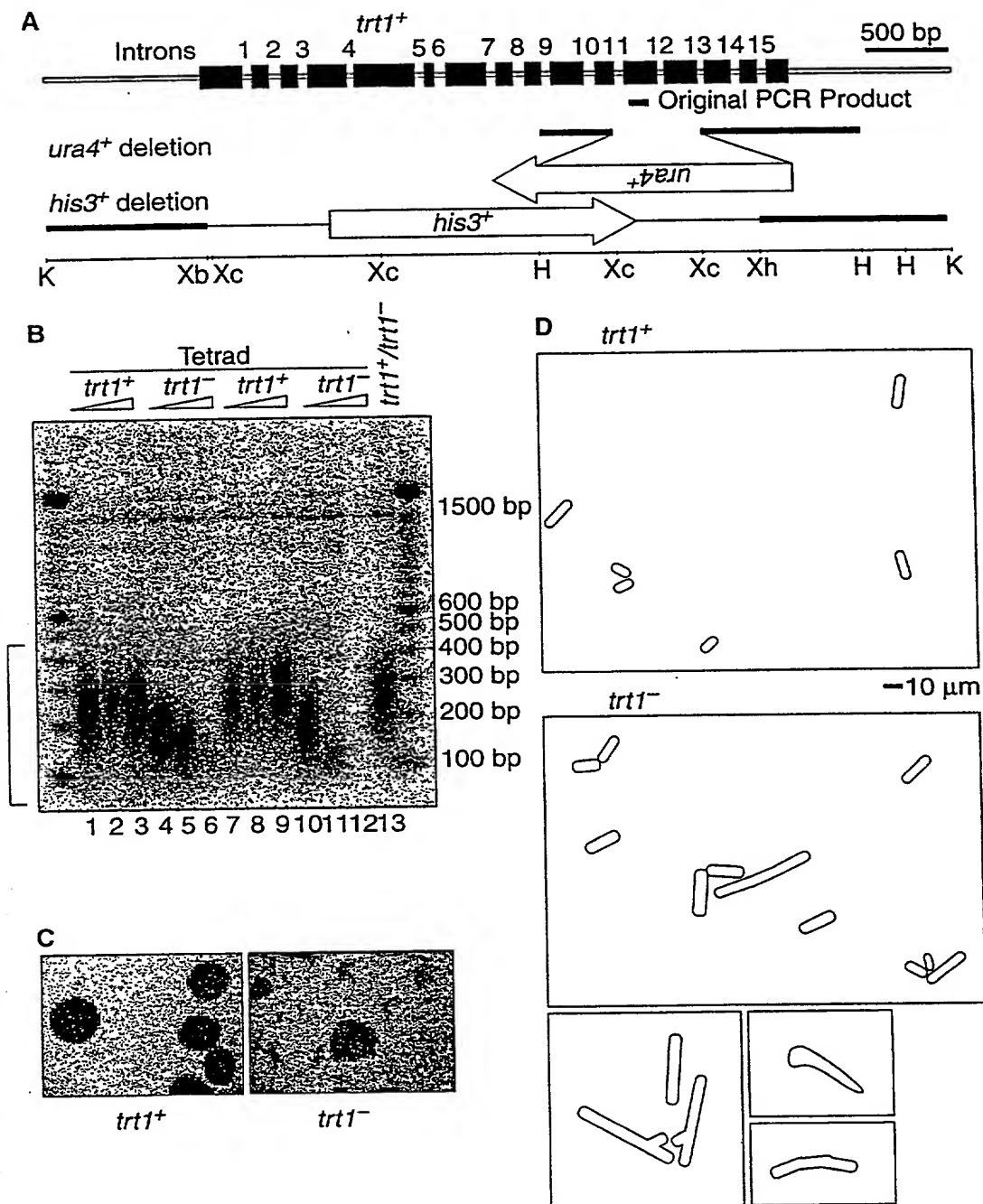


FIG. 22

33/34

gccaagttcctgcactggctgatgagtgtgtacgtcgtcgagctgctcaggtctttcttt
tatgtcacggagaccacgtttcaaaagaacaggctctttttctaccggaagagtgtctgg
agcaagttgcaaagcattggaatcagacagcacttgaagaggggtgcagctgcgggacgtg
tcggaagcagaggtcaggcagcatcgggaagccaggcccgcctgctgacgtccagactc
cgcttcatccccaaagcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
gccagaacgttccgcagagaaaaagagggccgagcgtctcacctcgaggggtgaaggcactg
ttcagcgtgctcaactacgagcgggcgcg

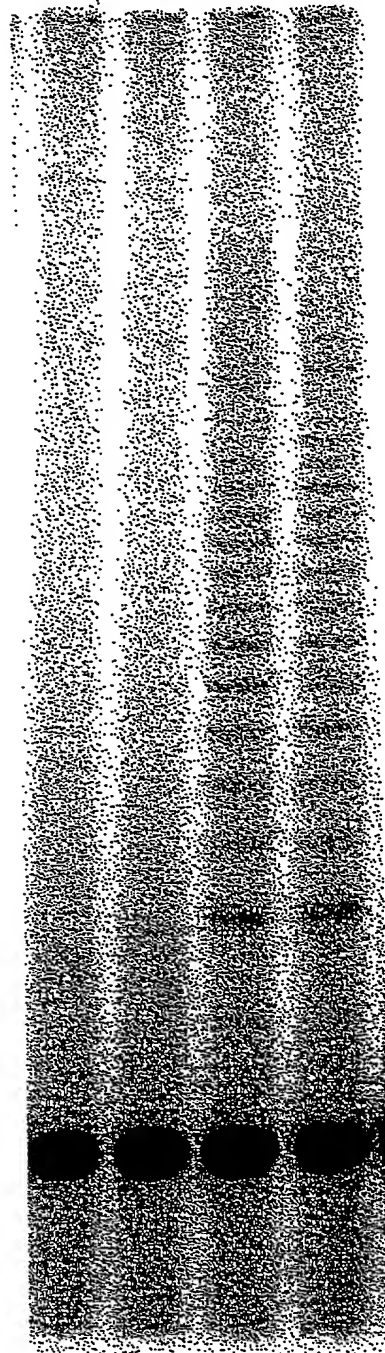
FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
ACCAGCCCCGTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC
AGGGGCAAGTC

FIG. 24

34/34

pBBS212 pGRN133



Internal Control

Approximate Cell No. 5,000 5,000 5,000 5,000

FIG. 25